Secondary Data: Valuable Resources for Extension Educators

Bo Beaulieu
Southern Rural Development Center
May 2007
What are Secondary Data?

- Information sources that already exist either in published or unpublished format.
- Data collected by someone else that are being “re-used” by others (such as Extension educators or university researchers).
- In contrast, “primary data” are those that you have collected.
- Secondary data can be either qualitative or quantitative in nature.
Understanding the Good, the Bad, and the Ugly about Secondary Data
The Good . . .

- They already exists!!
- Less expensive way to gather information
- Allows you to quickly get a handle on current and emerging issues
- Saves you the trouble of launching a more costly primary data collection effort
The Good . . .

- Saves you time (vs. primary data collection activities)
- May yield more accurate data than what you would obtain through primary data collection efforts (large vs. small samples)
- Can help to fine tune the focus of your primary data efforts, including your audience
The Bad . . . Or the Ugly!

- Inconsistencies in definitions (i.e., urban/rural, metro/nonmetro, farm size, family size, industry classifications)
- Data may be inaccurate or incomplete; biased
- Potential problems with “reliability” (i.e., sample size, response rate, questionnaire design)
- Data are usually only indirect measures of the issues you are interested in
- Data can be old
- Interpretation of the data must be done with caution (i.e., estimating impacts are difficult to determine directly; more likely to be indirect)
A Video Example of Confusion Regarding Data Definitions
Key questions you should ask

- What is the source of the data?
- Does it cover the correct geographical location?
- Does it provide data on the audience I’m interested in?
- Does it deal with the issue/topic I want to focus on?
- Does it represent current data?
- Are the data across time the same?
- Are definitions of the variables I’m interested in the same over time?
It’s Relevance to Extension

Can be used to:
- Understand the context for your educational programs
- Assess needs of your stakeholders
- Shape your evaluation indicators
- Monitor impacts
Some Good Sources of Data

- Federal Government
- Regional Organizations
- State Agencies
- Local Government
- Others
  - Trade associations
  - Private sector
Key Federal Data Resources

- **U.S. Census Bureau**
  - *Main Portal*
    - [http://www.census.gov/](http://www.census.gov/)
  - *Census Bureau A-Z Subjects*
  - *Census of Agriculture*
  - *State and Metropolitan Area Data Book*
    - [http://www.census.gov/compendia/smadb/](http://www.census.gov/compendia/smadb/)
  - *American Community Survey*
    - [http://www.census.gov/acs/www/](http://www.census.gov/acs/www/)
Key Federal Data Resources

- Economic Research Service

- Bureau of Economic Analysis
  - http://www.bea.gov/regional/reis/
Regional Data Resources

- Southern Region Information Toolkit (SRIT)
  - http://srit.srdc.msstate.edu/
State Level Data: Florida

- **Florida Statistical Abstract**
  - [http://www.bebr.ufl.edu/publications/Order_Series.shtml](http://www.bebr.ufl.edu/publications/Order_Series.shtml)

- **Office of Economic and Demographic Research (The Florida Legislature)**
  - [http://edr.state.fl.us/](http://edr.state.fl.us/)

- **Florida’s Health (Vital Statistics)**

- **Workforce Innovation**
  - [http://www.floridajobs.org/](http://www.floridajobs.org/)
Secondary Data Variables Of Relevance to Extension Programming

**POPULATION**
- Population Size
- Population Composition
- Population Distribution
- Migration Patterns

**EDUCATION**
- Attainment
- School Enrollment
- Dropout Status
- Performance Assessments

**ECONOMIC**
- Employment Status
- Income and Earnings
- Poverty Status
- Businesses/Firms
- Labor Force Composition
- Current and Future Jobs

**SOCIAL**
- Health and Nutrition Status
- Health Care Resources
- Crime Rates
- Housing
- Food Assistance Enrollment
- Child Care Access/Enrollment

**LOCAL GOVERNMENT**
- Revenues
- Expenditures
Okay, I’ve Found those Data. Now What?

Data analysis options:

- **Cross-sectional** -- look at data at one point in time
- **Comparative** -- examine the data in your county relative to other counties of interest
- **Longitudinal** – focus on how the data change over a longer time period
What to look for . . .

- Conditions that the data describe
- The direction of change
- The intensity of change
- How your county/community compares to other similar counties/communities
- The overall picture that the data paint about your county/community
Looking at Some Examples

Using Data to Paint a Picture
Cross-Sectional Data: Examining Local Conditions at One Point in Time
Educational Status Lower Among Rural Minorities

Nonmetro educational attainment by race and ethnicity, 2000

Percent of adults 25 and older

Hispanics may be of any race. "American Indian" and "Black" include Hispanics.
Poverty Status Among Minorities

Poverty rates by race and ethnicity, 2002
Nonmetro Blacks, Native Americans, and Hispanics have highest poverty rates

Percent poor

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>Nonmetro</th>
<th>Metro</th>
</tr>
</thead>
<tbody>
<tr>
<td>White, non-Hispanic</td>
<td>11.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>33.2</td>
<td>22.7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>26.7</td>
<td>21.4</td>
</tr>
<tr>
<td>Native American</td>
<td>34.6</td>
<td>19.2</td>
</tr>
<tr>
<td>Multiracial</td>
<td>19.5</td>
<td>12.4</td>
</tr>
</tbody>
</table>
Percent of Workers Employed in Creative Occupations in 2000

- Metro: 26.8% Creative, 73.2% Non-Creative
- Micro: 17.7% Creative, 82.3% Non-Creative
- Noncore: 14.8% Creative, 85.2% Non-Creative
Longitudinal Assessment: Examining Data Over Time
Nonmetro population change by county, 2000-05

Source: Prepared by ERS using Census Bureau 2005 population estimates.
Components of nonmetro population change by Census region, 2000-05

Number of residents (1,000)

Source: Calculated by ERS using Census Bureau 2005 population estimates.
Nonmetro population change by race/ethnicity and age, 2000-05

Source: Calculated by ERS using data from the U.S. Census Bureau.
In 1990, the Hispanic population was concentrated in the West, particularly the Southwest.

Source: William Kandel, ERS
By 2000, the Hispanics were settling in the nonmetro Midwest and Southeast.

Source: William Kandel, ERS

<table>
<thead>
<tr>
<th>Year</th>
<th>Goods Producing</th>
<th>Service Producing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1969</td>
<td>42.5%</td>
<td>57.5%</td>
</tr>
<tr>
<td>1990</td>
<td>33.8%</td>
<td>66.2%</td>
</tr>
<tr>
<td>2000</td>
<td>30.5%</td>
<td>69.5%</td>
</tr>
<tr>
<td>2004</td>
<td>27.3%</td>
<td>72.7%</td>
</tr>
</tbody>
</table>

Source: Bureau of Economic Analysis, Regional Economic Information Systems
Nonfarm Proprietor Jobs and Income as a Percent of All Jobs and Income in Nonmetro Areas, 1969-2000

Source: Goetz, 2005
Hands-On Activities: Putting the Data Puzzle Together